DOCUMENT RESUME

ED 245 795 PS 014 308

AUTHOR McNamara, Thomas C.
TITLE Charting the Course of Implementation.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jan 81
CONTRACT NIE-P-80-0184
NOTE 62p.; Paper prepared at the request of the National Institute of Education: Follow Through Planning Strand Two: Search for Future Through Approaches; Research/Methodological Issues.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Early Childhood Education; Models; *Program Evaluation; *Program Implementation; Socialization
IDENTIFIERS Judicial Evaluation Approach; *Project Follow Through; Sponsors; *Theoretic Orientation

ABSTRACT

The purpose of this paper is, generally, to discuss issues related to the evaluation of implementation and, specifically, to provide practical advice about Follow Through implementation and its evaluation. First, the discussion explores philosophical and theoretical assumptions regarding (1) the nature of implementation and evaluation; (2) realities of implementation as a unique form of socialization; and (3) approaches to evaluation that are compatible with the specified realities of implementation. Non-mechanistic, non-reductionistic philosophical assumptions and consonant theoretical positions are favored. Next, a discussion of realities of implementation provides a perspective on Follow Through evaluation efforts, analyzes implementation as socialization, and offers recommendations concerning the future of Follow Through. The concluding discussion of evaluation offers a concrete proposal for the evaluation of implementation. While many compatible evaluation stategies are listed, the judicial evaluation model is seen as the single approach most compatible with implementation realities, and a combined documenting and judicial evaluation approach is recommended. Briefly addressed are the inferential dimension of educational measures and the importance of Follow Through model sponsors. Appended is a rationale for a national Follow Through database that specifies needed categories of information and evaluation/research strategies. Also appended is a list of elements of child data and site descriptor computer files to be maintained at each Follow Through location. (RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED245795

PS 014308

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CHARTING THE COURSE OF IMPLEMENTATION

Thomas C. McNamara
School District of Philadelphia

Paper Prepared
At the Request of the
National Institute of Education:
Follow Through Planning, Strand Two;
Search for Future Follow Through Approaches;
Research/Methodological Issues
NIE-P-80-0184

January, 1981

CHARTING THE COURSE OF IMPLEMENTATION

I. INTRODUCTION

By placing this paper's discussion of implementation within the framework of "charting a course," it is intended that the focus be placed from the very outset on a few essential themes analogous to charting situations. These themes have to do with convictions like the following: (a) innovative program implementation in education is in large part a matter of being able to find one's way through a complex series of mazes — including the interactions of persons — surrounding the real world of the schools, (b) the pathways through these mazes are in reality merely passable routes between the interweaving motions of several layers of systems; routes that are only really open to experience and understanding, and (c) the whole journey through this system-maze is best captured by thorough, alert documentation.

To further set the stage for the discussion, and even more graphically try to convey the appropriateness of the notion of "charting" for dealing with this issue, it is perhaps useful to picture the whole innovation/implementation process as having some basic similarities with the attempt to produce sorely needed rain by seeding clouds with silver iodide. The seeding operation represents a carefully weighed decision to intervene in the continuous, massive movement of weather systems as they roll across the earth's surface. The decision to seed, furthermore, is made when rain is considered absolutely essential to remedy a

drought situation, and only after careful study beforehand, and follow-up tracking afterwards, of the – synoptic charts – which identify and locate the complex patterns of high and low pressure systems, temperature and humidity levels, effects of the prevailing winds at various altitudes, etc., which make up what we call weather.

The main points to be made, of course, in using this analogy, are that educational interventions, and study of their implementation, are no less a matter of deciding that some critical need calls for disrupting the otherwise steady flow of education's complex patterns, and then, on the basis of careful pre, during, and post charting routines, introducing the projected remedy at some fairly precise point, in some well-selected unit of education's huge, intricate system.

It might be further noted, in terms of appropriateness for implementation issues, that the weather analogy is likewise a clear reminder of at least two other points: (a) the primitive state of our knowledge of intervention agents in both areas – cloud-seeding is at best a crude and awkward process, with a lot of hit and miss characteristics, and our know-how with regard to educational interventions, despite all recent attempts, cannot honestly be considered to be much beyond the level of rainmaking expertise, and (b) the notorious fallibility of weather predictions, with educational predicting a not much more accurate counterpart. On the limping side of the analogy, it is of course, crucial to point out that the inanimate elements comprising the complicated web of interactions found in weather systems are far outdone in complexity by the self-knowing, abstract thinking beings we have to deal with in the educational sphere.

All of the above is not said by way of promoting discouragement, but only by way of setting what this paper considers a realistic background for taking up the charges assigned to it. These charges are essentially to discuss issues related to evaluating implementation, and to provide some practical considerations regarding implementation, and its evaluation as part of Strand Two (Search for Future Approaches) of NIE's efforts to provide redirection for Follow Through. The paper's response to these charges, keeping in mind the complex systems theme, and the weather analogue, introduced earlier, will be in the form of three major sections organized as follows:

1. Philosophical/Theoretical Assumptions Underlying this Discussion.
   Philosophical presuppositions very clearly affect one's outlook on the nature of implementation and evaluation, and consequently one's approach to evaluating implementation. This paper favors non-mechanistic, non-reductionistic philosophical assumptions – and theoretical positions consonant with this stance – because they appear to be more consistently meaningful in attempting to understand everyday educational reality.

2. Implementation Realities.
   A close look at implementation phenomena in the real world of education leads one to conclude that it is a unique form of socialization related to restructuring the perceptions, roles and actions of educating individuals functioning within at least four plausible modes of existing organizational structure (system), while allowing that mutual adaptation of innovation/existing system is always at least minimally involved.

3. Evaluation Compatible with Implementation Realities; Recommendations.

Not only is evaluation felt to be in need of a non-reductionistic base to be able to effectively engage in educational inquiry in general, but this is especially held to be the case in attempting to handle implementation phenomena with their continuously evolving process characteristics, i.e., complex human changes occurring against a backdrop of existing complex human organizations/systems.

The evaluation approaches that rely primarily on straightforward, comprehensive documentation, and those developed in analogy with judicial/quasi-judicial processes appear to be particularly appropriate for dealing with implementation because, on the one hand, they give prominence to thorough, continuous description, which is presently our best means of keeping track of the continuously shifting configurations of implementation "in vivo," and, on the other hand, feature, non-mechanical measures, i.e., well-reasoned judgments, to cope with educational reality, which is perhaps never more fluid, and less susceptible to fixed measures than when implementation is the focus.

Given the above premises, i.e., having the responsibility of accounting for the implementation process, with its primary characteristics of complex human changes being introduced into existing complex human systems, a sensible, logical strategy would seem to be a joint, interdependent use of comprehensive

documentation and quasi-judicial reviews/hearings, with the former serving as an ongoing feeder system for the latter, which would occur at regular intervals to synthesize and provide direction for each succeeding stage of implementation, and/or yield recommendations regarding the advisability of continuing any present attempt in its existing form.

II. PHILOSOPHICAL AND THEORETICAL ASSUMPTIONS REGARDING THE NATURE OF IMPLEMENTATION AND EVALUATION.

The paper makes several philosophical and theoretical assumptions about the unique capabilities of human beings, the participants in implementation, and, by implication, about the nature of implementation, which, in this conception necessarily involves introducing change in interaction with these unique capabilities:

(1) Human beings are persons, with characteristics that distinguish them qualitatively from other living things; among a number of essential characteristics are special cognitive capabilities that allow humans to be active knowers who know that they know, i.e., who are capable of thinking, of conscious planning and rule formulation based on the ability to engage in internal structuring and modifying of concepts.

(2) Given the above, the process of innovation/implementation necessarily implies, at least to some extent, having the persons involved change their thinking as regards a particular area of the educational process.

(3) Furthermore, the persons participating in innovation/implementation must be dealt with not only as separate individuals, but as already committed participants – again, to at least some degree – in an existing organizational structure (system), which defines current educational reality, in varying degrees, for these individuals.

The paper also makes assumptions about the nature of educational inquiry/educational evaluation, and hence about the nature of evaluation as applied to implementation. Again these assumptions rely on non-reductionistic views on the nature of human knowledge and educational phenomena.

A. Nature of Human Beings/Human Knowing

In making the above assumptions the paper is clearly siding with non-positivistic, non-associationistic views of humans and human knowing. Positivism, which has tended to govern, either formally or informally, the thinking of educational psychologists for most of this century, has always stood for an all-encompassing reductionism, i.e., there is nothing radically different about the way human beings know and learn, as distinct from other living things. Furthermore, in this view, especially in the form of

Behaviorism, it is also commonly held that relatively few, relatively simple laws relating neuronal excitation to response characteristics will suffice to explain what are usually considered unique human capabilities. This is particularly the case if these simple elements can be viewed as gradually combining via, again relatively few, relatively simple laws regulating their progressively more complex association.

The position endorsed by this paper does not consider Positivism or Associationism tenable philosophical positions in the face of everyday reality, perhaps especially everyday school reality. The mechanistic, reductionistic principles on which they rely are simply lacking in adequacy in the face of what are most characteristically human activities, particularly as regards human capability to function as purposive, self-knowing persons. With Wittrock (1974) and McNamara (1980) the paper believes it is time to recognize that we have been too long under the sway of Positivistic convictions, and to move on to a more balanced outlook on human beings and the ways in which they know. At the same time it is important to recognize the contributions of Positivism and Associationism, especially their preoccupation with clear thinking and specification, and avoidance of tendencies to rely on convenient, catch-all phrases that have not been subjected to rigorous analysis. It is important also to acknowledge the fact that there are undeniably some limited aspects of human behavior that can be adequately explained in mechanistic, reductionistic

terms, but these aspects do not comprise the full range of human capabilities, and, furthermore, it is only positions that can account for this full range which can legitimately, because knowingly, treat humans temporarily as if purely mechanical beings. This temporary process is a directly intended abstraction involving conscious disregard of the full complex of human capability for a time in order to treat humans in, e. g., purely quantitative terms to describe certain limited aspects of human phenomena.

B. Nature of Innovation/Implementation

1. Individuals Must Undergo Conceptual Changes.

There are some clear cut implications for conceptualizing the nature of educational innovation/implementation in making the above assumptions. The most fundamental implication is that introducing an innovation, and then seeing to its implementation, clearly involves conceptual changes in some sense on the part of participants. Human beings in this view, i.e., persons who know that they know and actively monitor and regulate this knowing process, especially when these human beings are themselves engaged in the daily process of teaching others, i.e., introducing conceptual change in systematic fashion to students, cannot be expected to abrogate this special human capability and respond in purely mechanistic terms when they themselves are the participants in an educational change process.

The thinking of Strike and Posner (1977) seems especially pertinent for further understanding the nature of the conceptual

change to be undergone by individuals according to the above views. and even more pertinent for understanding conceptual change in the context of curriculum organization, which must, by definition, underlie most proposed innovations/implementations. They support the following views of conceptual change — patterned after Toulmin's account of scientific procedure — as one of the most realistic:

> Conceptual change is evolutionary, not revolutionary... An evolutionary model of change allows us to evaluate any particular conceptual change against an enduring background of rational standards and explanatory ideals. Conceptual change is not a matter of judging between comprehensive views but rather of judging concepts in relation to other concepts, with respect to their adequacy in interpreting problems and fulfilling explanatory ideals... Concepts adapt to an intellectual environment much as organisms adapt to a biological environment.
>
> (Strike and Posner, 1977, p.111)

In subscribing to this outlook on conceptual change, Strike and Posner simultaneously endorse non-mechanistic positions regarding concept formation, cognitive structure, and thinking, proposing, in fact, that, as regards curriculum organization, "cognitive status and processes can be regarded as the 'embodiments' of the kinds of logical and conceptual features which characterize organized bodies of propositions" (op.cit., p. 136).

Similar convictions about the necessity for conceptual change in implementation have characterized the work of G. E. Hall and his colleagues at the Research and Development Center for Teacher Education at the University of Texas at Austin. Loucks and Pratt

(1979) describe the application of this group's "Concerns-Based Adoption Model" in the Jefferson County, Colorado schools. This approach makes the following four assumptions:

1. Change is a process, not an event.
2. Change is accomplished by individuals, not institutions.
3. Change is a highly personal experience.
4. Change entails developmental growth in both feelings about and skills in using new programs.

(Loucks and Pratt, 1979, p. 213)

McLaughlin (1978) has also concluded from study of both organizational change projects as well as general change – agent efforts that the primary focus in innovation and implementation must be on individual participants, and that mutual adaptation is essential. "The evidence we have seen strongly suggests that the developmental process of mutual adaptation is the best way to ensure that change efforts are not superficial, trivial, or transitory" (McLaughlin, 1978, p. 31). Mann (1978c) has extended this conclusion to the proposal that innovations take the form of "user-driven systems" in which individual teachers are given special incentive pay to raise the scores of targeted student groups.

2. The Individuals to Be Changed are Part of an Organization to be Changed

Though the above comments stress that individuals must change in innovations, and, therefore, imply that this paper is in clear-cut agreement with the Concerns-Based tenet, cited earlier, that "Change is accomplished by individuals,"

there is simultaneous reluctance to accept the second part of that tenet, i.e., "Change is not accomplished by institutions," without some distinctions. It is felt that until the day arrives – and perhaps it would be a good thing if it does – when a proposal like Mann's, also cited above, regarding direct incentives to teachers becomes a reality, (and perhaps even then, only in reduced form), it is only realistic to recognize that innovations/implementations presently involve individual teachers qua members of organizations (institutions), and that this organizational membership dimension must be taken into account in conceptualizing all that is important to implementation reality. It is in this context that Elmore's (1978) insights – coupled with the realistic approach of symbolic interactionism, to be discussed later – seem to offer a useful means of handling the organizational dimension. He has concluded, along with many others, that the widespread record of unsuccessful innovations is principally attributable to failures in implementation, and then draws the further conclusion that "implementation failures are failures of organization." Understanding organizations, therefore, is crucial in Elmore's thinking, if implemention is to be handled more successfully:

> I will develop four organizational models representing what I see as the major schools of thought that can be brought to bear on the implementation problem. The systems management model is my attempt to capture the organizational assumptions underlying the mainstream, rationalist tradition of policy analysis... The bureaucratic process model represents the sociological view

> of organizations, updated to include recent research by students of 'street-level bureaucracy'... The organizational development model represents a relatively recent combination of sociological and psychological theory that focuses on the conflict between the needs of individuals and the demands of organizational life. Finally, the conflict and bargaining model represents a tradition of inquiry that addresses the problem of how people with divergent interests coalesce around a common task.
>
> (Elmore, 1978, pp. 188-189)

After extensive elaboration of the models, Elmore reaffirms their usefulness for "certain commonsense explanations as to why social programs fail," and, in addition, feels that not only is there presently no conceivable way to reconcile these four different viewpoints, but this diversity allows different aspects of implementation problems to be highlighted and solutions formulated accordingly.

Kirst and Jung (1980) see the type of analysis proposed by Elmore as part of a second stage of theory development, following on an initial, descriptive, short-run case study phase, in attempting to explain the workings of implementation. They stress the timeliness of capitalizing on the advances made in stage two studies like Elmore's and perhaps ushering in a third phase by focusing on a longitudinal case study approach. They illustrate the approach by following the course of Title I implementation over 13 years and noting the changing influence of various interest groups on interpretations of implementation, as well as the continuous increase in the power of USOE officials to impose strict interpretations of compliance. Farrar, De Sanctis and Cohen (1980) would also argue for the importance of a longitudinal

perspective from a slightly different point of view, proposing an evolutionary connatation for implementation consonant with Strike and Posner's (1977) statements above about the evolutionary nature of conceptual change:

> Implementation is not the carrying out of a formulated policy but part of its evolution. And in that evolution — in certain circumstances or with certain programs — a multitude of local dispositions and actions are more important than the dispositions and actions of federal agencies.
> (Farrar, De Sanctis and Cohen, 1980, p.167)

It is perhaps useful to conclude these initial considerations on the nature of innovation/implementation by referring to Mitchell and Spady's (1978) article on the necessity of broader change movements, (but with implications for all innovations), such as the recent push for "competency based education," having to be viewed in an organizational context also, though in this case the organizational context has to do with outcome expectations that analysis would associate with four alternate bases of education, i. e., competency, development, social integration, and social responsibility.

C. Nature of Educational Inquiry/Evaluation

In the earlier part of this discussion, when the focus was on basic aspects of human nature, human knowing, and conceptual change, much has already been said that would be applicable here, i.e., the same rejection of reductionistic and mechanistic principles would be repeated. It only remains to summarize some statements by a number of major writers who have dealt more explicity with educational inquiry and educational evaluation, while simultaneously stressing the inadequacies of positivistic and similar assumptions.

First of all it is important that the nature of educational phenomena be clearly established under this viewpoint. Gowin (1972) makes a very precise statement:

> ...a most significant fact about educational phenomena: They are man-made (artifactual), not natural. They are therefore not likely to yield laws and other modes of invariance such as the natural sciences report in that domain. Whatever regularities researchers are to find in educational phenomena will have been determined by human beings in a social context. (Gowin, 1972, pp. 9-10)

Sarason (1978) reminds us that in social matters, and, therefore, equally so in education there are no "once-and-for-all solutions in the scientific sense"; essentially the same problems have to be solved over and over again. Scriven (1980) makes much the same comment, while proposing that educational inquiry "shed once and for all the sense of dependence on, of derivation from, of application via the social and behavioral sciences." For Scriven:

> Language is a better model of complex human behavior than the motion of small particles or large planets. We understand language; we understand many unpredictable and unprecedented linguistic events; we often understand why language is used. Yet we know no exact or statistical rules or laws of language. There are patterns in it; but not the precise patterns of planets or the probability distributions of particles. These are the patterns we must learn to find and describe in the study of teaching and administration and learning.
> (Scriven, 1980, p. 30)

House (1978) provided a taxonomy of major evaluation models, relating each to the philosophy of Liberalism; his major work on the "Logic of Evaluative Argument" (1977), however, seems to reveal his

preferences for models with non-positivistic bases, stressing, e.g., the continuous importance of "qualitative argument." In a very recent statement (House, 1980), he goes so far as to say that explanatory frameworks shift with changing social conditions:

> ...social science perspectives generally, describe or explain a social world which is itself constantly changing. At one time, a technological perspective may better explain the social reality, but at a later time the political perspective may do so. For example, as society becomes increasingly fragmented, political and cultural perspectives seem more salient as explanatory frameworks. (House, 1980, p. 3)

Guba (1980) and Sanders (1980) both argue convincingly against positivistic assumptions. Guba stresses the need for naturalistic versus conventional inquiry, while allowing that the "quantitative postures" of the conventional approaches are just as acceptable in naturalistic investigations, but under different basic assumptions. Sanders proposes a "developmental research" approach that recognizes the sui genesis nature and complexity of educational reality.

Strike (1972) stressed the need for Verstehen in educational inquiry:

> Man, for Weber, is a purposeful, rational, and value-seeking thing. Man has a subjective life. However, man's behavior can be accounted for scientifically by the simple expedient of incorporating man's subjective life into an account of his behavior. (Strike, 1972, p. 45)

In a statement conceptualizing the practical science nature of educational inquiry (Strike, 1979), he emphasized that education

is unavoidably engaged in the business of "normative and ends – means propositions," as well as "factual" and "theoretical" propositions.

Stake (Page and Stake, 1979) feels evaluation has nothing to fear from being "more subjective," and more recently (Stake, 1980) makes the point that it is an important role for researchers (and evaluators) to serve as facilitators for a "user's naturalistic generalizations," i.e., those "derived from tacit, personal, experiential learnings."

The above is only a brief summary of the assumptions/conclusions that will guide the discussion of evaluating implementation in the concluding section. As a prelude to that section. however, it seems best to further expand on some of implementation's realities, since evaluation of this phenomenon has to be particularly aware of and sensitive to all the continuously evolving aspects of implementation.

III. IMPLEMENTATION REALITIES

The preceding consideration of innovation/implementation was on a broader, philosophical/theoretical plane to provide the setting for this section which attempts to make the issue more concrete and amenable to explicit evaluation activities, as well as to prepare the way for some explicit recommendations regarding sponsor models/ implementation. The section will be divided into three main parts; in the first the attempt will be made to put Follow Through implementation efforts into perspective by applying Elmore's (1978) models;

in the second part attempting to analyze implementation as socialization, using Fullan's (1980) insights on this subject, as well as his delineation of implementation factors, and in the third part formulating some recommendations regarding the future of Follow Through based on these implementation considerations.

A. Follow Through Implementation in Perspective

House (1979) took the point of view that Follow Through was a child of its time in many respects, not the least of which was the fact that the Systems Analysis approach (a favorite of USOE and other federal agencies at the time), gradually took more and more solid hold of both program direction and the national evaluation during the program's first ten years. The Systems Analysis approach is one form of Elmore's (1978) Systems Management Model (Model I) and almost every point he makes in describing the characteristics of this model provide an equally accurate description of many of the implementation tribulations of Follow Through – and also what seems to have been some of the prime assumptions about the nature of implementation in USOE's mind, the sponsor's mind, and sites' mind, as can be seen in, e.g., Haney's (1977a, 1977b), Kennedy's (1977, 1978), Wisler, Burns and Iwamoto's (1978), and Hodges et al.'s (1980) discussions of this issue. This was perhaps inevitable, given the fact that, in Fullan and Pomfret's (1977) terms, the "fidelity" – in contrast to the "mutual adaptation" – conception of implementation was apparently taken for granted – or was the exclusive choice, perhaps, in some cases – by seemingly all concerned. It should also be noted that during most, if not all

of the early years of Follow Through, the spirit – if not the letter – of positivism still exercised considerable influence over convictions regarding educational research and evaluation, not to mention over educational theories that kept substantial ties to early behaviorism, and, as already mentioned, over prevailing convictions about implementation, even if only in once – removed form, e.g., via Systems Analysis. It seems to be fair to all concerned to say that Elmore's (1978) Model I (Systems Management) and Model II (Bureaucratic Process) go far in explaining practically every aspect of the direct program implementation convictions and operations that have existed throughout most of Follow Through up to this point. It seems, in fact, not to be overstating the case to see some analogies between a widespread outlook on what should constitute Follow Through Model implementation, at least during the program's early years, and what Goodenough (1966) and other anthropologists have referred to as "cargo cult" expectations. As will be discussed below, there was some relationship between Follow Through and Elmore's Model IV (the Conflict and Bargaining Model), but apart from large city situations, where powerful Unions had to be dealt with, Model IV came into play primarily in relation to survival rather than implementation per se.

Elmore (1978) has indicated that within the framework of Model I (Systems Management) implementation failures are logically attributable to "bad management," i.e., not specifying goals clearly,

not assigning/monitoring responsibility carefully enough, etc. Furthermore, there is a normative dynamic to the model which dictates that, if implementation is failing, it is not the fault of the Systems Management Model — since things ought to be this way — so the only recourse is to restructure the situation to fit the model. (Elmore feels the model works best in private business settings where the kind of control the model calls for can, for the most part, be exercised as a matter of course.) For Model II, the Bureaucratic Process Model, implementation problems are seen as principally a matter of failing to influence existing organizational routines, so that as a result, they interfere with the innovation's operations. This model's normative side does not have the force of Model I and simply consists in stressing the importance of somehow influencing the routines that might be problemmatic. The history of Follow Through and its handling of the "implementation problem" seems well captured within the boundries of these two models, for the most part, and principally within Model I.

It is in order at this point to explain the comment made above about Model IV (Conflict and Bargaining) and Follow Through. This connection seems best explained — again, apart from large city Union problems — in saying that Model IV appeared to enter the scene not very much in relation to implementation per se, but rather in relation to survival. Parents from sites continuously and vigorously lobbied for the program's survival, but in itself

implementation was not the foremost issue in the lobbying effort. Model IV could also be said to have characterized Sponsor-USOE relationships after a time, but again survival, rather than implementation appears to have been the uppermost consideration — though perhaps implementation figured here also, but under the form of what Haney (1977a) calls the "Penelope Ploy," i.e., the continuous postponement of the effort at final implementation specifications, etc., so that effectiveness considerations — whether consciously or unconsciously — would not have to be dealt with in clear-cut terms.

To return to the central theme of the present discussion, however, i.e., Follow Through implementation efforts as covered by Elmore's (1978) Models I and II, it is important to note that in taking the above position this paper is not trying to detract from what Follow Through has accomplished in the implementation sphere. With Haney (1977a/1977b), Wisler, Burns and Iwamoto (1978), and Hodges et al. (1980), the paper has great appreciation and admiration for the program's being able to obtain any successes at all in the face of having to operate without a clear definition of purpose; "unplanned" Planned Variation, in many respects; constant shifting of the ground rules, so to speak; etc. There is likewise great appreciation for what Follow Through Sponsors have accomplished, despite all odds, in pioneering a new third party role, as documented by not only the writers mentioned above, but also, e.g., by Kennedy (1977, 1978), Nero and Associates (1975),

and Rath (1976). But, as is evident from the preceding section on philosophical/theoretical considerations, the paper would naturally have to be more interested in Follow Through relationships with Elmore's (1978) Model III (Organizational Development), than his other models; those relationships, however, seem tenuous at best, if not altogether nonexistent. The paper logically would have to see broad normative influence of Model III as the ideal situation.

Model III posits that failures in implementation "are not the result of poor management control or the persistence of bureaucratic routines, but arise out of a lack of consensus and commitment among implementers" (Elmore, 1978, p. 209). In Elmore's thinking, even the idea of "mutual adaptation" does not completely explain all that is meant by "organizational development," since:

> The real significance...is that it effectively turns the entire implementation process on its head. It reverses what we instinctively regard as the 'normal' flow of policy, from top to bottom...The result is that, in terms of the effective structure of organizations, the process of initiating and implementing new policy actually begins at the bottom and ends at the top...The role of those at the top of the system, then, is necessarily residual...
>
> If one accepts this view, the important business of implementation consists not of developing progressively more sophisticated techniques for managing subordinates' behavior but enhancing the self-starting capacity of the smallest unit. (Elmore, 1978, pp. 211-212)

It is one of the main convictions of this paper that not only is it time for Follow Through to move consistently in this direction, but that the program will be stunted in its natural development if

it does not — if it is even allowed to continue to exist as before. More will be said on this point in the recommendation discussion below. Before making the recommendation, it seems useful to further expand on implementation realities by pursuing the concept of implementation as socialization with Fullan (1980).

B. Implementation as Socialization

According to Fullan (1980), implementation is a socialization process with varying degrees of interaction among personnel, and between personnel and materials, strategies and overall philosophical commitment. There seems to be little doubt that framing implementation in a socialization context goes a long way toward concretely characterizing what has usually taken place in implementation efforts, and that this approach can complement Elmore's (1978) insights especially where Model I (Systems Management) and Model II (Bureaucratic Process) seem to be the ruling influences. Implementation as socialization seems less applicable to Model IV (Conflict and Bargaining), and to Model III (Organizational Development), though, of course, relevant to some degree here also. Symbolic interactionist viewpoints like those of Becker (1968) and culture and personality perspectives like those of Spindler (1968), with their emphases on discovering/maintaining identities within socialization situations might well be able to reconcile any apparent discrepancies between at least Model III and "socialization." Mischel's (1979) recent work on "cognitive economics" might also provide some light on how disparate

individual/societal (implementational) demands might be harmonized, and Hall's (1980) work on implementation configurations would be useful as well.

In any case, as already noted, treating implementation as socialization certainly seems to have descriptive utility, along with Elmore's models, for providing enlightenment as to what has happened in the past, especially the Follow Through past, and these are important considerations in filling out this discussion of implementation realities. Fullan (1980) also sees two principal sets of factors as affecting implementation/socialization. These factors — and subfactors — are as follows:

1. Characteristics of the change effort
    a. The adoption decision
    b. Clarity
    c. Complexity
    d. Implementation strategies
    e. External/internal relationships
2. Characteristics of the Institutional Setting
    a. History of innovative attempts
    b. Role of central administrators
    c. Role of principal
    d. Organizational characteristics
    e. Community characteristics

In the following, in an effort to provide further grounding in reality for implementation issues, each of these subfactors will receive brief comment, either with reference to statements in the professional literature or from historical Follow Through events:

1a. The adoption decision. Haney (1977a, 1977b) and others remind us that little time and little planning were associated with initial Follow Through decisions. Mann (1978b) feels most adoption decisions are unfortunately made on opportunistic bases. Evaluation of NDN (1977) findings give some precise recommendations about , e.g., personal involvement at all levels in the adoption decision.

1b. Clarity. The two most "structured" models in Follow Through appear to have been the most clearly specified. Apart from their theoretical base which would stress precise formulation, one wonders whether or not Cronbach's (1970) distinction between "maximum" versus "typical" behavior might be applicable here also, which would imply that the structured approaches were able to be more precise — after the fashion of achievement tests — because they seemed to continually focus on eliciting "maximum" behavior, in contrast to the more unstructured approaches — with parallels in personality measures, and their problems, etc., — which were more concerned with developing "typical" behavior.

1c. Complexity. Fullan (1980) tends to feel there is an inverse relationship between clarity and complexity. McLaughlin (1978), on the other hand, in her study of organizational change projects, found implementation was not only not impeded, but actually enhanced by complexity. It is perhaps worth noting that Snow (1980) has recommendations for "demystifying" the concept of "learner-control," which has represented quite a complexity for

some of the unstructured models. Child Development considerations could also perhaps be made less complex by tending to, e.g., Elkind's recommendations on using Piaget as a thinker rather than an educator (cited by Albin, 1980).

1d. Implementation strategies. Fullan (1980) feels it may be counterproductive to try to explain everything at once, and that it may be especially important to postpone considerations of the innovation's philosophical bases until participants are more receptive. ASCD Update (1980) notes that there is widespread teacher discontent with support in their curriculum development needs. Perhaps Kessen's (1979) recommendation on seeing child development in a cultural context, and Sroufe's (1979) insights on the coherence of individual development should be the kind of background context provided for most innovation.

1e. External/internal relationships. It is perhaps ironic that experience with the various Follow Through Models sometimes led to the conclusion that the more structured the model, the more flexible it was in accommodating local needs and circumstances, and the contrary for the more unstructured, which sometimes, appeared to have either taken a very defensive position, or else indicated that they had nothing to prove.

Sponsor internal stability and therefore continuity of model development were crucial factors in Follow Through.

2a. History of innovative attempts. Fullan notes that Sarason's "prehistory" concept is of great importance here. In large cities, especially, it would seem that Sponsors had to overcome a good deal of cynicism or a "show me" attitude because of all the "panaceas" to which these school systems had been subjected since the 50's. Mann (1978a) notes the primary and complexity of the sites in all cases; a complexity that an unfortunate "prehistory" could compound.

2b. Role of central administrators. The innovations of modern education very often stand or fall with administrators' approval /commitment, which is apparently often tied to their own opportunities/ambitions. Some administrators, on the other hand, might resort to a "power strategy" (Patterson and Czajkowski, 1979).

2c. Role of the principal. Fullan sees the principal's supportive role as absolutely essential; he unfortunately has concluded that 50-60% of the nation's principals are not instructional leaders, but administrators only, and this fact often hinders implementation. From another point of view, it may turn out that some few individual principals may be able to set up a school environment that is perfectly in keeping with all the control factors of Elmore's Model I (Systems Management).

2d. Organizational characteristics. Dealing with Unions is a critical matter, for large cities especially;there should be concerted efforts at all levels, federal, state, local – to

either neutralize any obstacles stemming from existing Union agreements – perhaps by concluding a special, separate agreement – or by including Union members as essential participants in the original implementation decisions so that this element of "ownership" generates cooperation throughout the project's duration. Other factors that have had considerable effect on implementation have been, e.g. (1) strikes, especially long, bitter strikes, (2) periodic funding crises resulting in lowered morale, (3) school district reorganizations during the year, (4) lack of funding for adequate staff development, and (5) conflicts with other federal funding sources, such as Title I eligibility demands, conflicting program foci, etc.

2e. Community characteristics. An incident of some relevance here is the following: a model with an essential family outreach component was selected for an area of the city with crime-ridden high-rises; the location effectively helped stymie implementation; staff assigned home contact responsibilities were simply afraid to carry them out and avoided doing so whenever they could.

Bridge (1978) makes some useful recommendations regarding parent involvement, e.g., when and when not to strive for group consensus.

C. Recommendations for Future Follow Through Implementation Efforts

It has already been anticipated in earlier statements that the main tenor of this recommendation will be that Follow Through Models (Sponsors) move very steadily toward not just a "mutual adaptation" mode, but toward Elmore's Model III approach, i.e., a highly

committed organizational development stance that focuses on developing "self-starting capacity" in each teacher. (The Follow Through Program in Philadelphia has begun moving in this direction on its own, first establishing local derivations of models, and combinations of models, and more recently initiating a "best elements" approach, selecting the best aspects of a number of models.) The principal implication here is that Sponsors should primarily become co-facilitators with site personnel in developing personalized models for each teacher — which in turn implies cross-sponsor cooperation, various forms of "hybrid-model" thinking, etc., in order to best serve each individual teacher in developing his/her highest possible level of teaching expertise. All that has been learned in functioning at the global model level can now be taken the next step forward in a plausible developmental sequence, i.e., personalizing, tailor-making the models in a very literal sense. There seems to be a very real mission for Follow Through in this conceptualization, one that has to do with close alliances with sites in working toward the effective development of highly skilled teachers in compensatory education through enabling model-grounded elements to become realized in unique, personal, "best-fit" models.

In this conceptualization, teacher commitment would be built in by definition, since the approach presupposes that sites would be funded to work with low-achieving schools through enlisting teacher support for further self-development, which in turn would involve Sponsor(s) expertise in helping sites accurately diagnose the needs

(expressed and observed) of each teacher so that self-motivated efforts result, geared toward developing a personalized model for each teacher. Each site would have available a panel representing a range of structured/unstructured sponsors (realistically, perhaps no more than a team of three or four sponsors would be needed) who could work cooperatively together and with local training staff in helping each teacher develop an effective, unique teaching style based on a special combination of best model elements. The approach, of course, still relies on some total program commonalities: small group instruction, multi-adult classrooms, parent involvement, etc.

The above recommendation makes one further presupposition about the possibility of close Sponsor cooperation, and that is that a flexible, eclectic position be taken vis-à-vis one another, so that whatever fundamental philosophical differences may at one time have been associated with particular models, these be put aside in the interests of developing effective approaches that represent the best form of rational cooperation in a world where no one theory/approach has proven to be unassailable.

In the evaluation section below a major research and evaluation role will be recommended for Sponsors as uniquely able and experienced investigators and, therefore, much in demand in moving the study of implementation forward and expanding on e.g., Snow's (1977) and Cronbach and Snow's (1977) urgings regarding aptitude – treatment interaction through special expertise in developing the teacher-treatment aspect in the form of personalized models.

IV. EVALUATION COMPATIBLE WITH IMPLEMENTATION REALITIES

This final section will attempt to bring together the points made above about implementation, both as regards philosophical/theoretical considerations and its more concrete realities, as illustrated in the last section, and what was said earlier about philosophical/theoretical premises in the evaluation area, intending, as a result, to produce a concrete proposal for the evaluation of implementation. The section will be divided into three parts: (a) a general consideration of evaluation strategies compatible with the philosophical/theoretical assumptions made earlier, (b) special focus on the single strategy considered most appropriate, and (c) a final recommendation regarding the evaluation of implementation.

A. General Consideration: Evaluation Strategies Compatible with Earlier Assumptions

1. Some Compatible Approaches

As a general introduction to the approaches to be listed below, it might be useful to give prominence to Goodlad's (1979) two recommendations on research into classroom management, since they reflect the basic orientation of many of the strategies listed:

> The first is that conclusion-oriented researchers interested in understanding classroom phenomena get involved in the "stuff" of practice — in naturalistic, ethnographic inquiries...The second

> suggestion is that decision-oriented researchers interested primarily in improvement should collaborate with teachers in inquiries in which assumptions and purposes are shared.
> (Goodlad, 1979, p. 407)

The compatible strategies include, among many other candidates: Blumer's (1970) symbolic interactionist outlook with emphasis on direct "exploration" and "inspection" of the "empirical world"; Carini's (1975) "descriptive research"; Bronfenbrenner's (1976) "ecological experiment" approach; Busses, Chittenden, and Amarel's (1978)" collaborative research"; Bateson's (1980) insistence on the primacy of description versus our usual predilection for sanctioning "facts" only after processing in the "magical" realm of explanatory mathematical models; Guba's (1978) general endorsement of naturalistic versus conventional inquiry; Stake's (1978) recommendations on case studies and naturalistic generalizations; Eisner's (1979) proposals for effective use of an "educational criticism" strategy; Weiss' (1980) suggestions for turning our attention to "nonconventional" outcomes of schooling, and their assessment, and Kean and McNamara's (1978) "gradual refinement" approach, possibly with Apple's (1980) additional recommendations.

2. A Brief Reflection on Measurement

In leading up to the final recommendation, it will serve a useful purpose to briefly reflect on the nature of educational measurement. Jones (1971) reminds us:

> Some measurements can be made quite simply and directly... Other attributes, however, can be estimated only from their effects... The inference about an attribute from its effects involves either an assumption or a demonstration of a relation between effect and attribute.
> (Jones, 1971, p. 337)

We often forget that all of our educational measures involve inferences from effects. The article by the Laboratory of Comparative Human Cognition (1979) quotes Goodenough to much the same effect, noting that our educational and psychological measures are not "measuring devices, properly speaking. They are sampling devices." Cronbach (1971), and Messick (1975, 1980) have always stressed the prime importance of construct validity precisely because judgment/inference is at the heart of our measurement process. Cronbach (1970) sees construct validity as the "long — continued interplay between observation, reasoning and imagination," and Petrinovich (1979) likens the entire process of scientific inquiry to the search for construct validation. The main point is that perhaps we have been going in the wrong direction in our attempts to better our instruments; it seems reasonable to argue that bettering our measures is essentially a matter of refining our judgments, making our inevitable inferences much more sharply focused and competent.

3. The Need for Model Sponsors in Evaluation

Evaluation, especially the evaluation and study of implementation is one of the most important functions Follow Through Sponsors can provide. Not only are they now uniquely

experienced in the struggle to deal with implementation of their models, but they also have all the advantages, e.g., Gallagher (1979) attributes to research centers, i.e., situations allowing long-term commitment to instrument development, sustained and organized concentration on major problem dimensions and their interrelationships, and the possibility of being in a position to discover the unexpected, important insight. Follow Through Sponsors are an irreplaceable third party unit in this paper's conceptualization of Follow Through's future because of their dual capabilities as unique co-facilitators in the development of personalized models, as already mentioned, and as uniquely equipped evaluators of the complex implementation process.

B. The Single Most Compatible Approach: Evaluation as Judicial Hearing Analogue

Wolf (1979, 1980) is, as is well known, one of the chief proponents of the Judicial Evaluation Method (JEM):

> the law, as a metaphor, offers many important concepts (fact-finding, adversarial proceedings, cross-examination, evidentiary rules and procedures, structured deliberations, etc.)that when adapted to evaluation efforts add certain dimensionality lacking in more conventional forms of social inquiry... The ultimate evidence, then, which guides deliberation and judgment includes not only the "facts," but a wide variety of perceptions, opinions, biases, and speculations, all within a context of values and beliefs. Oftentimes the more subjective forms of evidence help put the facts into proper perspective.
> (Wolf, 1979, pp. 20-21)

He is aware of the many criticisms leveled against the process, but but replies that most of these come from misunderstanding the intent,

and suffer from taking the "judicial" in literal, rather than metaphorical form. One of the most recent criticisms is that by Worthen and Rogers (1980). They feel the approach – which they label "adversary," an identifier continually avoided by Wolf – is not useful for formative decisions and is "most appropriate for large, controversial programs which had a variety of interested audiences" (Worthen and Rogers, 1980, p. 540). It is important to remember, however, that the judicial "metaphor" is the main point, and that with Thurston (1978), for instance, several other judicial forms than the jury trial are available and may be more useful on occasion, e.g., the appellate court model and the administrative hearing model. It is this paper's position that the implementation process being the evolutionary, multi-faceted phenomenon that it is, involving conceptual changes, reconciling of organizational roles, etc., this is exactly where fixed measures are least able to cope, and the judgmental base of all measures becomes most important, so why not formalize this importance through use of some judicial or quasi-judicial method? These methods can range from some of the peer review processes suggested by Apling (1979) to the fullest kind of parallel with a judicial process. The basic procedure is the crucial consideration, and can be made to involve enough stages and safeguards to insure against premature "final judgments," while at the same time, keeping decisions out in the open frequently enough to avoid postponing those critical go/no go decisions when they should

be made. Walker (1977) indicates that Schaffarzick:

> suggested that greater efforts at systematic consideration of curriculum change may be more likely when the community context is more aroused and polarized on an issue than on more routine issues because the arguments and evidence resulting from systematic exploration of the issue may be more effective persuaders in such a climate.
>
> (Walker, 1977, p. 295)

The judicial analogue approach has immediate relevance and usefulness in this same vein, because it allows the teacher participants in implementation to have a forum for settling what may be very real problems in dealing with the required changes of a specific innovation.

C. Basic Conclusion/Recommendation: Combined Documenting and Judicial Evaluation Approaches

The paper began with a systems theme, and it is appropriate that it end on that theme. (Note that "systems" here has nothing in common with Systems Analysis (Systems Management.) Messick (1972) and Messick and Barrows (1972) have stressed the need for comprehensive documentation and regard for the fact that we are continually dealing with varying forms of systems in education:

> To recapitulate briefly, it looks as if the nature of the thing to be changed, the educational arena, is not only a complex system composed of many constituents but a complex culture comprising multiple roles and pluralistic values, and that if we are to understand the functioning of that system — or to change it — we must take into account the interplay of those roles and values in supporting (or subverting) system regularities.
>
> (Messick, 1972, p. 79)

McNamara (1980) recommended that comprehensive, computerized networks of basic school experience data (for longitudinal tracking) — that would translate here into implementation data — be developed by sites that would serve as a "skeletal" structure to be continually flushed out by ongoing case studies, participant observation, and collaborative research. The Research and Evaluation Committee of the Follow Through SCAN Forward Planning Task Force* has recommended a national Follow Through Data Base (Appendix). Hanson and Schutz (1978) report on a "product systems" approach, and Cooley and Leinhardt (1980) have recently concluded an instuctional dimensions study within a similar framework. These are the kind of documenting systems, which, when combined with various types of judicial hearings at regular intervals — perhaps 3 times a year formally, but once a month in more informal fashion, e.g., in the form of reviews, etc. — would appear to be a sensible way of capturing and constantly contributing to the improvement of the important, evolutionary process of implementation. This is felt to be particularly true of Follow Through implementation, whether this remains a matter of only some mutual adaptation on the part of sponsor-site, or, what is considered the ideal situation in this paper, represents a steady move toward directly personalized models.

*Members: M. Wang, T. Gennari, J. Lumbley, S. Fox, J. Rogers, D. Revicki, E. Ellett and T. McNamara

REFERENCES

Albin, R. S. David Elkind: Going Beyond Piaget. APA Monitor, 11(11) 4-5, Nov. 1980

Apling, R. Examining Evaluation Strategies for Improving Early Childhood Title I Program Practice. AERA Symposium Presentation: The Development of Evaluation Policy for Early Childhood Title I Programs, San Francisco, Calif., April, 1979

Apple, M. W. Putting Schools as Socio-Economic Institutions Back Into Educational Evaluation: A Response to Kean and McNamara. Unpublished Paper, University of Wisconsin – Madison, 1980

ASCD Study Finds Teachers' Curriculum Improvement Needs Poorly Met. ASCD Update, 22 (4) 1,6, Sept., 1980

Bateson, G. Mind and Nature, N. Y., N. Y. E. P. Dutton, 1979

Becker, H. S. The Self and Adult Socialization. In E. Norbeck, D. Price-Williams, W. M. McCord (Eds.) The Study of Personality: An Interdisciplinary Appraisal, N.Y., N.Y.: Holt, Rinehart and Winston, 1968, pp. 194-208

Blumer, H. Methodological Principles of Empirical Science. In N. K. Denzin (Ed.) Sociological Methods, Chicago, Ill.: Aldine, 1970, pp. 20-39

Bridge, R. G. Parent Participation in School Innovations. In D. Mann (Ed.) Making Change Happen?, N.Y., N.Y.: Teachers College Press, 1978, pp.101-119

Bronfenbrenner, U. The Experimental Ecology of Education Educational Research, 5(9) 5-15, Oct., 1976

Bussis, A.M., Chittenden, E. A. and Amarel, M. Collaborative Research. Paper presented by the ETS Early Education Research Group, Third Annual Conference on Research in the Arts and Aesthetic Education, Aspen, Colorado, June, 1978

Carini, P. F. — Observation and Description: An Alternative Methodology for the Investigation of Human Phenomena, Grand Forks, N. D.: University of North Dakota, Feb., 1975

Cooley, W. W. and Leinhard G. — The Instructional Dimensions Study. Educational Evaluation and Policy Analysis, 2 (1), 7-25, Jan.-Feb., 1980

Cronbach, L. J. — Essentials of Psychological Testing (3rd ed.), N.Y., N. Y.: Harper and Row, 1970

Cronbach, L. J. — Test Validation. In R.L. Thorndike (Ed.) Educational Measurement (2nd ed.), Washington, D.C.: American Council on Education, 1971, pp. 443-507

Cronbach, L. J. and Snow, R. E. — Aptitudes and Instructional Methods, N.Y., N.Y.: Wiley, 1977

Eisner, E. W. — The Use of Qualitative Forms of Evaluation for Improving Educational Practice Educational Evaluation and Policy Analysis, 1 (6). 11-19, Nov.-Dec., 1979

Elmore, R. F. — Organizational Models of Social Program Implementation. In D. Mann (Ed.) Making Change Happen?, N.Y., N.Y.: Teachers College Press, 1978, pp. 185-223

Evaluation of the National Diffusion Network, Executive Summary Menlo Park, Calif.: Stanford Research Institute, Oct., 1977 DHEW Contract 300750260

Farrar, E. DeSanctis, J. E. and Cohen, D. K. The Lawn Party: The Evolution of Federal Programs in Local Settings. Phi Delta Kappan, 62 (3) 167-171, Nov., 1980

Fullan, M. and Pomfret, A. — Research on Curriculum and Instruction Implementation. Review of Educational Research, 47 (2), 335-397, Spring, 1977

Fullan, M. — Research on Implementation. Presentation at the Collaborative Planning Meeting for 1980-1981 Follow Through, Southwest Educational Development Laboratory, Austin, Texas June, 1980

Gallagher, J. J. — Research Centers and Social Policy. American Psychologist, 34(10) 997, Oct., 1979

Goodenough, W. H. — Cooperation in Change, N. Y., N. Y.: John Wiley & Sons, Inc., Science Editions, 1966

Goodlad, J. I. — Perspectives on Theory, Research and Practice, In D. L. Duke (Ed.) Classroom Management, 78th NSSE Yearbook, Part II, Chicago, Ill.: University of Chicago Press, 1979, pp. 391-412

Gowin, D. B. — Is Educational Research Distinctive? In L. G. Thomas (Ed.) Philosophical Redirection of Educational Research, 71st NSSE Yearbook, Part I, Chicago, Ill.: University of Chicago Press, 1972, pp.9-25

Guba, E. G. — Toward a Methodology of Naturalistic Inquiry in Educational Evaluation. CSE Monograph Series in Evaluation #8, Los Angeles, Calif.: UCLA, 1978

Guba, E. G. — Naturalistic and Conventional Inquiry. AERA Symposium Presentation: Considerations for Educational Inquiry in the 1980's, Boston, Mass., April, 1980

Hall, G. — The Concept of Innovation Configuration. Presentation at the Collaborative Planning Meeting for 1980-1981 Follow Through, Southwest Educational Development Laboratory Austin, Texas, June, 1980

Haney, W. A. — A Technical History of the National Follow Through Evaluation, The Follow Through Planned Variation Experiment, Vol. V, Cambridge, Mass.: The Huron Institute, August, 1977, (a)

Haney, W. A. — The Follow Through Experiment: Summary of an Analysis of Major Evaluation Reports. Curriculum Inquiry, 7,227-257, Fall, 1977(b)

Hanson, R. A. and Schutz, R. E. — A New Look at Schooling Effects from Programmatic Research and Development. In D. Mann (Ed.) Making Change Happen? N.Y., N.Y.: Teachers College Press, 1978, pp. 120-149

Hodges, W., Branden, A., Feldman, R., Follins, J., Love, J., Sheehan, R., Lumbley, J., Osborn, J., Rentfrow, R. K., Houston, J., and Lee, C. Follow Through: Forces for Change in the Primary Schools, Ypsilanti, Mich.: The High/Scope Press, 1980

House, E. R. The Logic of Evaluative Argument. CSE Monograph Series in Evaluation, #7, Los Angeles, Calif.: UCLA, 1977

House, E. R. Assumptions Underlying Evaluation Models. Educational Researcher, 7(3) 4-12 March, 1978

House, E. R. The Objectivity, Fairness, and Justice of Federal Evaluation Policy as Reflected in the Follow Through Evaluation. Educational Evaluation and Policy Analysis, 1(1) 28-42, Jan. - Feb., 1979

House, E. R. Mapping Social Disconsensus onto Social Theory. AERA Symposium Presentation: Considerations for Educational Inquiry in the 1980's. Boston, Mass., April, 1980

Jones, L. V. The Nature of Measurement. In R. L. Thorndike (Ed.) Educational Measurement (2nd. ed.), Wash., D. C.: American Council on Education, 1971, pp. 335-355

Kean, M. H. and McNamara, T. C. The Gradual Refinement Approach to Large Scale Compensatory Education Programs. Evaluation and Program Planning, 1,161-162, 1978

Kennedy, M. M. The Follow Through Program. Curriculum Inquiry, 7, 183-208, Fall, 1977

Kennedy, M. M. Findings from the Follow Through Planned Variation Study. Educational Researcher, 7, 3-11, June, 1978

Kessen, W. The American Child and Other Cultural Inventions. American Psychologist, 34(10) 815-820, Oct. 1979

Kirst, M. and Jung, R. The Utility of a Longitudinal Approach in Assessing Implementation: A Thirteen Year View of Title I, ESEA. Educational Evaluation and Policy Analysis 2(5), 17-34, Sept. - Oct., 1980

Laboratory of Comparative Human Cognition. Cross-Cultural Psychology's Challenges to Our Ideas of Children and Development. American Psychologist 34(10) 827-833, Oct., 1979

Loucks, S. and Pratt. H. A Concerns – Based Approach to Curriculum Change. Educational Leadership, 37(3) 212-215, Dec., 1979

Mann, D. Introduction. In D. Mann (Ed.) Making Change Happen?, N.Y., N.Y.: Teachers College Press, 1978, pp. XI-XXVI.(a)

Mann, D. The Politics of Training Teachers in Schools. In D. Mann (Ed.) Making Change Happen?, N.Y., N.Y.: Teachers College Press, 1978, pp. 3-18.(b)

Mann, D. The User-Driven System and a Modest Proposal. In D. Mann (Ed.) Making Change Happen?, N.Y.,N.Y.: Teachers College Press, 1978, pp. 285-307.(c)

McLaughlin, M. W. Implementation as Mutual Adaptation. Change in Classroom Organization. In D. Mann (Ed.) Making Change Happen?, N.Y., N.Y.: Teachers College Press, 1978, pp. 19-31

McNamara, T. C. Ongoing Documenting Systems. AERA Symposium Presentation: Considerations for Educational Inquiry in the 1980's, Boston, Mass., April, 1980

Messick, S. Research Methodology for Educational Change. In Educational Change: Implications for Measurement, Proceedings of the 1971 Invitational Conference on Testing Problems, Princeton, N.J.: Educational Testing Service, 1972, pp. 69-81

Messick, S. The Standard Problem: Meaning and Values in Measurement and Evaluation. American Psychologist, 30(10) 955-966, Oct. 1975

Messick, S. Test Validity and the Ethics of Assessment. American Psychologist, 35(11) 1012-1027, Nov., 1980

Messick, S. and Barrows, T. S. Strategies for Research and Evaluation in Early Childhood Education. In Early Childhood Education, 71st NSSE Yearbook, Part II, Chicago, Ill.: University of Chicago Press, 1972, pp. 261-290

Mischel, W. On the Interface of Cognition and Personality. American Psychologist, 34(19) 740-754, Sept., 1979

Mitchell, D. E. and Spady, W. G. Organizational Contexts for Implementing Outcome Based Education. Educational Researcher 7(7) 9-17, July/August, 1978

Nero and Associates A Description of Follow Through Sponsor Implementation Processes, Portland, Oregon: Nero and Associates, Inc., April, 1975. ERIC #ED 128 926.

Page, E. B. and Stake, R. E. Counterpoint: Should Educational Evaluation Be More Objective or More Subjective? Educational Evaluation and Policy Analysis, 1(1) 45-47, Jan. - Feb., 1979

Patterson, J. L. and Czajkowski, T. J. Implementation: Neglected Phase in Curriculum Change. Educational Leadership, 37(3) 204-207, December, 1979

Petrinovich, L. Probabilistic Functionalism: A Conception of Research Method. American Psychologist, 34(5), 373-390, May, 1979

Rath, S. W., Sposito, P., Chapin, R. and Newton, K. Follow Through: A Story of Educational Change, Portland, Oregon: Nero and Associates, Inc., June, 1976. ERIC #ED 129-418

Sanders, D. P. Educational Inquiry as Developmental Research. AERA Symposium Presentation: Considerations for Educational Inquiry in the 1980's, Boston, Mass., April, 1980

Sarason, S. B. The Nature of Problem Solving in Social Action. American Psychologist, 33(4) 370-380, April, 1978

Scriven, M. — Self-Referent Research (II) Educational Researcher, 9(6) 11-18, 30, June, 1980

Snow, R. E. — Learning and Individual Differences. In L. S. Shulman (Ed.) Review of Research in Education #4, Itasca, Ill.: F. E. Peacock Publishers, 1977, pp. 50-105

Snow, R. E. — Aptitude, Learner Control, and Adaptive Instruction. Educational Psychologist, 15 (3), 151-158, Fall, 1980

Spindler, G. D. — Psychocultural Adaptation. In E. Norbeck, D. Price-Williams, and W. M. McCord (Eds.) The Study of Personality: An Interdisciplinary Appraisal, N. Y., N. Y.: Holt, Rinehart and Winston, 1968, pp. 326-347

Sroufe, L. A. — The Coherence of Individual Development: Early Care, Attachment, and Subsequent Developmental Issues. American Psychologist, 34(10) 834-841, Oct., 1979

Stake, R. E. — The Case Study Method in Social Inquiry. Educational Researcher, 7(2) 5-8, Feb., 1978

Stake, R. E. — Generalizations. AERA Symposium Presentation: Considerations for Educational Inquiry in the 1980's, Boston, Mass., April, 1980

Strike, K. A. — Explaining and Understanding: The Impact of Science on Our Concept of Man. In L. G. Thomas (Ed.) Philosophical Redirection of Educational Research, 71st NSSE Yearbook, Part I, Chicago, Ill.: University of Chicago Press, 1972, pp. 26-46.

Strike, K. A. — An Epistemology of Practical Research. Educational Researcher, 8(1) 10-16, Jan., 1979

Strike, K. A. and Posner, G. J. — Epistemological Perspectives on Conceptions of Curriculum Organization and Learning. In L. S. Shulman (Ed.) Review of Research in Education #4, Itasca, Ill.: F. E. Peacock Publishers, 1977, pp. 106-141

Thurston, P. Revitalizing Adversary Evaluation: Deep Dark Deficits or Muddled Mistaken Musings. Educational Researcher, 7(7) 3-8, July/August, 1978

Walkser, D. F. Toward Comprehension of Curricular Realities. In L. S. Shulman (Ed) Review of Research in Education, #4 Itasca, Ill.: F. E. Peacock Publishers, 1977, pp. 268-308

Weiss, J. Assessing Nonconventional Outcomes of Schooling.: In D. C. Berliner (Ed.) Review of Research in Education, #8, Washington, D. C.: American Educational Research Association, 1980, pp. 405-454

Wisler, C. E., Burns, G. P. and Iwamoto, D. Follow Through Redux: A Response to the Critique by House, Glass, McLean, and Walker. Harvard Educational Review, 48(2) 171-185, 1978

Wittrock, M. C. Learning As a Generative Process. Educational Psychologist, 11(2) 87-95, 1974

Wolf, R. L. The Use of Judicial Evaluation Methods in the Formulation of Educational Policy. Educational Evaluation and Policy Analysis, 1(3) 19-28, May-June, 1979

Wolf, R. L. The Way I See It ... American Education Should Go On Trial. Educational Leadership, 37(7) 533-535, April, 1980

Worthen, B. R. and Rogers, W. T. Pitfalls and Potential of Adversary Evaluation. Educational Leadership, 37(7) 536-543, April, 1980

APPENDIX

Draft:
A National Follow Through
Data Base

Research and Evaluation Committee
SCAN Forward Planning Task Force
Spring, 1980

D R A F T

A NATIONAL
FOLLOW THROUGH DATA BASE
RATIONALE

Introduction

A new direction is set for Follow Through, to take effect in 1981. No longer is the program officially targeted for dismantling by the federal government; instead it is now sanctioned as a program with two different, though complementing missions. Its first mission essentially continues the original service goal of the program, though, as previously, the program even in this respect is not conceived in purely service terms. There is also now explicit emphasis on managing this uniquely comprehensive educational effort to a degree of effectiveness that makes this nominally "service" operation a de facto demonstration program. Eighty percent of the sites nationally are scheduled to be this kind of Service Project.

Under its second mission, Follow Through is now to engage in pilot Knowledge Production activities in 20% of its sites. In this respect the program takes on a specific research resource dimension that can simultaneously provide consistent improvement for its demonstration (Service) commitment to the education profession, and continual growth and new development in basic understanding of the teaching/learning phenomenon in early childhood.

Given this new, two-fold direction for Follow Through — and remembering its highly controversial history, both programmatically and in terms of evaluation, — it appears to be an important consideration to provide a mechanism

that will allow this multi-faceted program to be continuously documentable in all major respects, and at the same time to be transformed via this documentation into readily available configurations of variables that can be meaningfully researched and evaluated. A National Follow Through Data Base is proposed to serve as this essential mechanism.

The concept of the National Follow Through Data Base will be fully expanded in the sections below. By way of further clarification of why it is proposed, the immediately following comments are offered regarding the kinds of information that are needed for this program, as well as its research and evaluation needs.

Kinds of Information Needed

Follow Through is a highly complex program, and even more complex now in its "new look." First it is a "multidimensional" program. That is, it serves as a "natural resource," demonstration program (Service aspect), for early childhood development and education through four program components (education, parent involvement, staff development, and support services). Follow Through, furthermore, emphasizes not only service related to these four components, and planned variation, but their evaluation as well, and now in its Knowledge Production role also has a clear-cut research responsibility.

A second characteristic of Follow Through in its Service capacity is that it is still an experiment in planned variation. As a result, Follow Through continues to include a diversity of educational models. Although all are aimed at promoting early childhood development and education, the models differ in philosophical and theoretical orientations and in the specific

approaches they employ.

In addition to the workings of each of the four general program components, the contribution of each of the models represented in Follow Through must be recognized in any documentation/evaluation effort for the overall program. However, providing information only on the mode of operation and effectiveness of each model is insufficient. Personnel in the United States Dept. of Education must be able to present Congress with data about the overall impact of the program for use in deliberations regarding future authorizations. Thus, evaluation efforts must include an emphasis on the program as a whole in relation to each of the four components, as well as preserve a correct perspective on each sponsor's contribution within each component.

The multidimensionality and diversity of the Follow Through program suggest that at least two kinds of documenting/evaluative information are needed under its service aspect. First, there must be broad-based information that is obtained by all sponsors at each participating site. This information will be useful in assessing and refining the service that Follow Through provides, and then will in turn feed the Knowledge Production aspect by helping to identify potentially beneficial efforts in the areas of early childhood development and education, and, more specifically, in generating innovative evaluation/research strategies for studying the impact of a wide range of interventions in a variety of contexts.

A second kind of evaluative information required for the program's main function is information related to the needs and interests of particular sponsors and participating sites. This information is crucial in a national evaluation of Follow Through service, but at the same time it must be

recognized that its prime purpose is to contribute to the assessment and refinement of specific models in their service capacity, and to add to existing research knowledge regarding educational innovation (Knowledge Production dimension).

Specific information related to the needs and interests of particular sponsors and sites cannot be provided in this proposal. However, recommendations can be made regarding the kinds of information that should be gathered, or at least furnished, by the total program, i.e., across all sponsors. Five types of data are of particular importance; and would be of major concern in the construction of a National Follow Through Data Base:

a. Site characteristics

b. Program characteristics

c. Delivery system characteristics

d. Degree of implementation

e. Outcomes

Each type of information is briefly discussed below:

a. Site Characteristics

Previous studies have suggested that individual sites differ radically in terms of a variety of variables, e. g., student characteristics, staffing, school and community setting. Data should be gathered on these variables for possible use in explaining some of the outcomes that are achieved at particular sites. Variables that are found to be significantly related to outcomes can then not only remain the concern of evaluation, but also

be a focus of future research efforts.

b. Program Characteristics

At least two characteristics of the model programs designed and implemented by Follow Through sponsors are of particular importance. These characteristics are: objectives and critical features. Most of the programs have the same general objectives, but they differ in the importance they assign to each and in the specific approaches used to meet these objectives. Information on objectives and approaches can be used in interpreting evaluation results. As an example, the importance of low scores on measures of personal and social development would be minimized for a program that emphasizes basic skills. Instead, the program would be judged primarily on outcome data related to the basic skills objectives that it sought to attain.

c. Delivery System Characteristics

Model programs differ not only in their objectives and features, but also in the delivery systems they employ, e.g., number of sites, structure of delivery system, staff involved, etc.

A National Data Base should also take this type of program diversity into account and include the collection of data on characteristics of the delivery systems implemented by various Follow Through sponsors.

d. Degree of Implementation

The need for including implementation data has become increasingly

apparent in relation to Follow Through. There are at least three reasons why such information is essential. First, sponsors have to clearly define their models. These definitions should then become operational so that relative judgments about levels of model implementation across sites and across schools or classrooms (or even families, for some models) within sites can be made. Inability to produce implementation scales in some cases may be due to lack of clarity in the definition of the model. Such a lack of clarity probably foreshadows future difficulties.

Second, implementation scales are needed to provide information that permits the linking of outcome data to implementation levels. Some sponsors have already begun using implementation data in their evaluation designs. From the sponsors' perspective, such designs are superior to designs concerned only with outcomes, since they allow some separation of model characteristics from local site variations and, hence, represent truer tests of the model itself.

A third reason for good implementation data stems from an interest in implementation as an outcome itself. For some sponsors, changes in the educational process represent key outcomes of the sponsor intervention apart from any concern with products such as test scores. It appears important that interest in implementation as an outcome should not be restricted to these sponsors alone.

There are some further important aspects of implementation that should be noted at this point in view of a National Data Base effort:

In evaluations of sponsor delivery systems, implementation variables will be the most direct indices of delivery systems effectiveness. Such evaluations should be valued for formative purposes as well as externally if Follow Through is truly to live up to its potential as a "natural resource", both for demonstration and research purposes.

For within-sponsor comparisons, implementation ratings need to be reliable and comprehensive, but can be model-specific. For across-sponsor comparisons, there is a need for implementation judgments that have some common meaning for all sponsors involved, at least to the point of quantifying the meanings of "low," "moderate," and "high" levels of implementation. An approach to the development of implementation scales is proposed in the section on implementation below.

e. Outcomes

To date, an almost exclusive focus in determining the effectiveness of Follow Through has been student achievement. However, given the multidimensionality of the program, it is critical that data be gathered on other outcomes related to the education component (e.g., student autonomy, self-concept), as well as to the other three components of the national program. Examples of outcomes that might be considered include: quality of school environments, student/staff interactions, and community involvement. As noted earlier, model implementation may also be viewed as another outcome, as may some specific site characteristics, e.g., ability to transform family status.

Evaluation/Research Strategies Needed

Once decisions are made regarding the specific kinds of information that are required in order to meet the information needs of various stakeholders in the Follow Through program, evaluation/research strategies for collecting and analyzing this information must be identified. Two issues related to such strategies seem of particular significance. The first concerns the need for alternative strategies, the second with program change.

As indicated in the previous section, a wide variety of information is required in order to study and assess a program as complex as Follow Through. One large-scale experimental study such as was conducted previously cannot possibly provide the information required. In planning a National Data Base for the new Follow Through, alternative designs, measures, and analysis procedures should be identified and the most appropriate ones selected or developed.

In terms of alternative designs, any soundly constructed design that offers the opportunity of increasing existing knowledge should be considered. This approach, while limited by economic constraints, would open up Follow Through to much more innovative documentation, and research. These designs could contribute significantly to the entire field of educational investigation, as well as furnish information that would be beneficial to program stakeholders at the local, sponsor, and national levels.

Of particular interest at this time is a combination of quasi-experimental and descriptive research. Descriptive data cannot only serve as a complement to data gathered through experimental methods, but serve as useful investigative means in their own right. Descriptive data collected on classroom processes,

for example, can prove invaluable in generating research questions for field experiments aimed at refining program components. Further, these data are useful in accurately documenting the process and outcomes of particular programs under certain specified conditions.

The second major issue related to evaluation/research strategies concerns program change. During the course of any given year and particularly from year to year, if the past is any indication of the future, there will be changes in the Follow Through program at all levels. National objectives may be revised, objectives of particular model programs may be refined, and new delivery systems for implementing programs at participating sites may be developed. Thus, it is critical that evaluation/research strategies are capable of providing documentation of such changes. Further, if certain strategies are deemed inappropriate, given program changes, then these strategies must be modified or replaced by strategies that can more effectively meet program needs, and documentation that can handle this contingency should be available.

These strategies are based on the assumption that the "new look" Follow Through Program will build in adequate funding from year to year for research and evaluation in the service dimension as well as in the Knowledge Production area, and that there will be continual give-and-take in research and evaluations ideas and proposals between sponsors, sites, and the Dept. of Education.

D R A F T

A NATIONAL FOLLOW THROUGH DATA BASE

PRINCIPAL ELEMENTS OF THE FILE

The basic concept of a National Follow Through Data Base assumes a common computerized file creation effort across the program. This common effort would principally revolve around the development of two types of files to be maintained at each site: (1) a child data file and (2) a site descriptor file. In addition each sponsor might find it profitable to maintain a separate file system that contains key elements of the child data and site descriptor files in existence at each of the sponsor's sites; the sponsor systems should probably also include the kinds of delivery system characteristics outlined in Table 1 below:

I. CHILD DATA FILE ELEMENTS

A. Fixed Length Portion of File

1. ID - common ID approach across all sites — basically a numeric system with initial alphanumeric character to distinguish site.
2. NAME - fixed as well as phonetic spelling — the latter for better tracking later on, in case of use/disuse of initials, etc.
3. BIRTHDATE
4. RACE CODE
5. SEX CODE
6. SES INDICATORS - e.g., perhaps median income/% on welfare by site, geographic areas, etc.
7. PRE-K EXPERIENCE CODE - basically yes/no; perhaps type

8. ENTERING K/FIRST GRADE TEST DATA - essential pre-program data to be derived from common instrument/subtest across all sites.

B. Variable Length Portion of File

N.B. For each year after the child enters the program:

1. SCHOOL YEAR
2. SCHOOL CODE
3. GRADE
4. CLASSROOM IDENTIFIER
5. LENGTH OF PROGRAM EXPOSURE (also model indicator in multi-model sites) — e.g., in months
6. YEARLY ABSENCE RATES
7. SPECIAL EDUCATION/GRADE REPEAT INDICATORS
8. LOCAL SITE STANDARDIZED NORM-REFERENCED ACHIEVEMENT TEST SCORES (Possibly better managed by pre-post testing on a matrix-sampling basis using a common instrument across all sites)
9. SPECIAL SPONSOR TEST DATA
10. CRITERION-REFERENCED COGNITIVE/NON-COGNITIVE TEST DATA (TO BE DRAWN FROM A COMPREHENSIVE ITEM POOL TO BE ESTABLISHED ACROSS SPONSORS AS SOON AS POSSIBLE, e.g., BY END OF NEXT FIVE YEARS AT THE LATEST)

II. SITE DESCRIPTOR FILE

A. Fixed Length Portion of File

1. SITE CODE (As in CHILD DATA FILE)

2. BASIC PROGRAM DATA — e.g., # Schools, # Teachers, # Aides # Children, Funding Sources, etc.

B. Variable Length Portion of File

1. SCHOOL YEAR
2. CROSS-PROGRAM, DEGREE OF IMPLEMENTATION DATA BY SCHOOL, GRADE, CLASSROOM # — large representative sample each year; using "high," "medium," "low" indicators relative to ideal implementation
3. SPECIAL SPONSOR IMPLEMENTATION DATA
4. SPECIAL SITE IMPLEMENTATION DATA (OPTIONAL)
5. SUMMARY TEST DATA DERIVED FROM CHILD DATA FILE;
6. PARENT INVOLVEMENT INDICATORS — e.g., some basic data regarding # volunteers per family, # volunteer hours, # workshops attended, # other activities attended
7. STAFF DEVELOPMENT INDICATORS — e.g., some basic data regarding general staff experience, experience in program, # consultations received, # workshops attended, etc.
8. SUPPORTIVE SERVICE INDICATORS — e.g., some basic data regarding # medical/dental screenings, # treated; # psychological referrals, # treated; # families contacted, # in need of social services, # being helped through social services
9. SPECIAL "NATURALISTIC INQUIRY" INDICATORS — e.g., possibly special status (total site/specific aspect) codes derived from participant observation, in-depth interviews, questionnaires, etc., describing some

"emerging" program characteristics as it functions in the real world; maybe there should be some cross-program arrangements for continuous, representative efforts at participant observation

The site descriptor file could be further specified by taking into account the list of variables appearing in Table 2.

Table 1

Examples of Delivery System Characteristics

1. Number of sites served
2. Total number of people involved in delivery system at the sponsor level
   (a) number of full-time personnel
   (b) number of part-time personnel
   (c) number of people responsible for delivery of services to individual sites (person/site ratio)
3. Total number of people involved in delivery system at the site level
4. Organizational structure of delivery system at the sponsor level (e.g., types of personnel, types of direct and indirect service)
5. Organizational structure of delivery system at the site level
6. Time spent (in terms of delivery services) in contact with individual sites per year
7. Types of interactions (on-site interactions and sponsor-shop interactions) that take place as part of delivery system
8. Types of sponsor/site evaluative feedback processes utilized

Table 2

Examples of Site Characteristics

1. Students
   - (a) mobility
   - (b) age
   - (c) sex
   - (d) ethnicity
   - (e) participation in program (voluntary/involuntary)
   - (f) previous education (nursery school)

2. LEA staff
   - (a) age
   - (b) sex
   - (c) ethnicity
   - (d) experience
   - (e) education

3. School district/school
   - (a) geographic distribution
   - (b) # of years in FT program
   - (c) comprehensive services offered
   - (d) district organization/structure (relation of FT director to district organizational structure)
   - (e) administrative organization
   - (f) extent/use of local resources
   - (g) leakout to NFT

(h) distribution of FT classrooms vs. entire school district

(i) additional federal program participation with FT population

(j) proportion FT/district budgets

(k) distribution of sites (re: sponsor and sponsor location)

(l) sponsor/site "marriage" reasons (who, how, why)

(m) sponsor/site "divorce" reasons

(n) summer/after-school activities

(o) proportion of exceptional children in FT classrooms

(p) retention statistics

(q) desegregation plan (process used)

(r) union participation

(s) assignment of classrooms to Follow Through (voluntary/involuntary

4. Community

(a) rural/urban

(b) population

(c) SES